Retrospective and Prospective Case Studies to Accelerate the Pace of Chemical Risk Assessment

Katie Paul Friedman[1], Matthew Gagné[2], Tara Barton-Maclaren[2], John Bucher[3], Russell Thomas[3], Mike Rasenberg[4], Tomasz Sobanski[4]

[1]National Center for Computational Toxicology, US EPA, Research Triangle Park, NC USA; [2]Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa ON; [3]National Toxicology Program, NIEHS, Research Triangle Park, NC USA; [4]Computational Assessment Unit, European Chemicals Agency, Helsinki, Finland

www.epa.gov

1737/P113

March 12, 2019

Society of Toxicology Annual Meeting Baltimore, MD

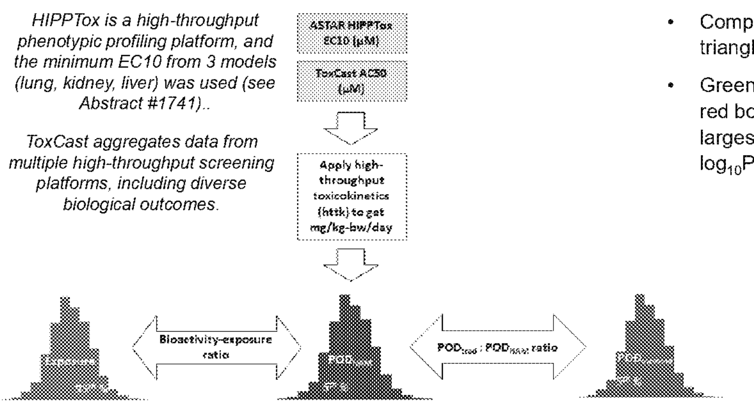
Katie Paul Friedman I paul-friedman.katie@epa.gov

Abstract

Use of high-throughput, *in vitro* bioactivity data in setting a point-of-departure (POD) has the potential to accelerate the pace of human health risk assessments by chemical prioritization. Advancement toward this goal requires confidence that *in vitro* bioactivity data, in conjunction with high-throughput toxicokinetic information, can be used to estimate administered equivalent doses at or below the PODs from traditional animal studies. Further, hazard and exposure predictions, combined as a bioactivity:exposure ratio (BER) for use in risk-based prioritization, should be evaluated. In this work we describe two efforts of the Accelerating the Pace of Chemical Risk Assessment initiative, a consortium of international regulatory scientists, both with the same primary objective: to elucidate whether a POD derived from *in vitro* bioactivity would be a conservative estimate of traditional POD estimates, and if the BER is a useful prioritization metric. In the first project, we describe the outcome of a retrospective case study of 448 chemicals with high-throughput predictions of bioactivity, reverse dosimetry, and exposure, as well as traditional hazard information. For 92% of these chemicals, a POD derived from new approach methodologies (POD_{NAM}) was a conservative prediction for the traditional POD ($POD_{traditional}$) value. High-throughput exposure predictions were greater than the POD_{NAM} for 26/448 chemicals, with BERs of less than zero, indicating higher priority for further investigation. The second, prospective study involves generation of NAM data for 200 chemicals to prioritize 20 chemicals for 90-day repeat dose testing in rats using a combination of the BER and bioactivity-based flags. Together these case studies enable regulatory scientists from different international contexts to develop efficient approaches for chemicals management, while possibly reducing the need for animal studies. This work demonstrates the feasibility, and continuing challenges, of using bioactivity and exposure NAMs in screening level safety assessment. *This abstract does not necessarily reflect ECHA, Health Canada, NTP, or U.S. EPA policy.*

Part I: Retrospective case study

Figure 1. Overall retrospective workflow

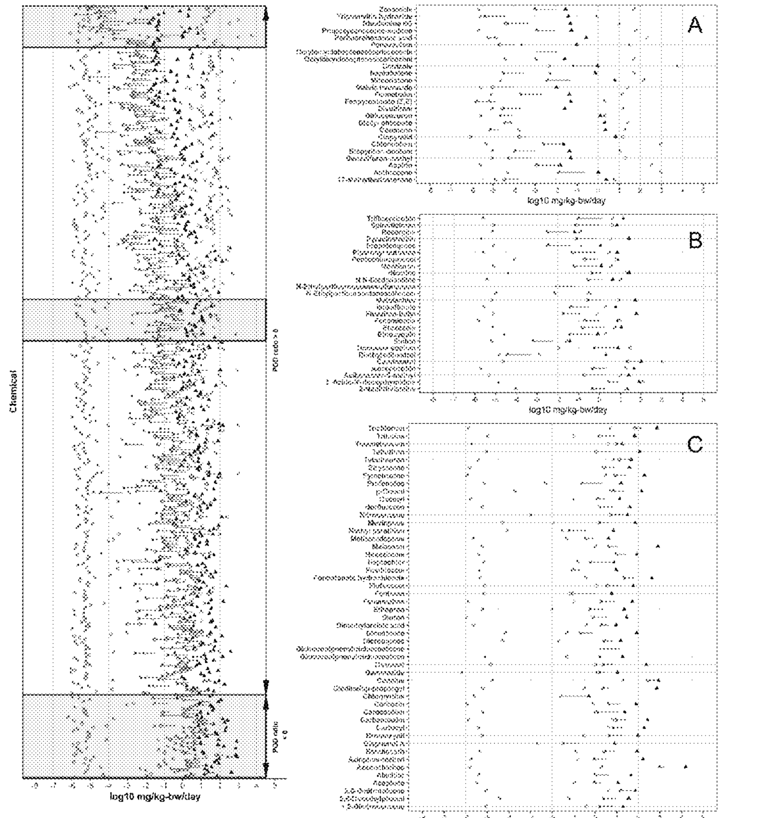


- 448 substances with exposure predictions (ExpoCast SEEM2 95th percentile for total US population), *in vitro* assay data, HTTK information, and *in vivo* hazard information.
- 50th and 95th percentile from the Monte Carlo simulation of inter-individual toxicokinetic variability were used to estimate administered equivalent doses (AEDs) for the minimum HIPPTox EC10 and the 5th percentile of credible ToxCast AC50 values for each substance.
- The minimum of either the ToxCast or HIPPTox-based AEDs were selected as the $POD_{NAM,50}$ or $POD_{NAM,95}$. The POD_{NAM} estimates were compared to the 5th percentile from the distribution of the $POD_{traditional}$ values obtained from multiple sources to obtain the log_{10}POD ratio.
- The log_{10}BER was obtained by comparing the POD_{NAM} estimates to exposure predictions. All values used for computation were in log_{10}-mg/kg-bw/day units.

$POD_{NAM,95}$ would have been conservative for screening and prioritization purposes when compared to $POD_{traditional}$ for 89% (400/448) of the substances.

Figure 2. Comparison of Exposure, POD_{NAM}, and $POD_{traditional}$

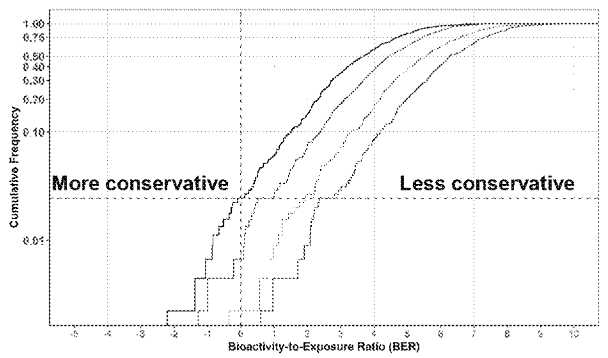
- Comparison of ExpoCast (SEEM2; gray circles), POD_{NAM} (green circles), maximum AED (black triangles), and $POD_{traditional}$ values (blue boxes) for 448 substances.
- Green line segment indicates the $POD_{NAM,95}$ to $POD_{NAM,50}$. Inset images A, B, and C correspond to the red boxes overlaid on the main plot. Image 3A provides a magnification on the substances with the largest log_{10}POD ratio values. Image 3B displays a sample of substances that approach the median log_{10}POD ratio. Image 3C includes all 48 substances for which the $POD_{NAM,95}$ > $POD_{traditional}$.



NAM-based approach informs reasonable, conservative screening and prioritization

Figure 3. Cumulative frequency of bioactivity-exposure ratio (BER)

- BER_{95} used 95th percentile from the credible interval to predict median total US population exposure (ExpoCast SEEM2);BER_{50} the 50th percentile.
- BER_{95} and BER_{50} values were calculated as the "95th%-ile" and "50th%-ile," using the $POD_{NAM,95}$ and $POD_{NAM,50}$, respectively.

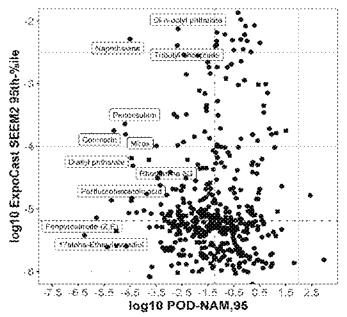


11 of 448 substances had a BER_{95}, 95th%-ile < 0

BER_{95}, 95th percentile did not prioritize an unreasonable number of substances; the BER selected reflects the level of conservatism and uncertainty considered within a screening assessment

Figure 4. Did exposure or bioactivity appear to drive the BER-based priority?

- Compared 95th percentile from the credible interval to predict total US population exposure (ExpoCast SEEM2) to the $POD_{NAM,95}$.
- Dashed lines indicate the median exposure and $POD_{NAM,95}$ estimates for the 448 substances in the case study.



In general for log_{10}BER < 0, the POD was relatively low. For certain substances the exposure estimates were relatively low.

Figure 5. Comparison of $POD_{NAM,95}$ and TTC

$POD_{NAM,95}$ > TTC for 87% of the substances (389/448); the two values do not appear correlated

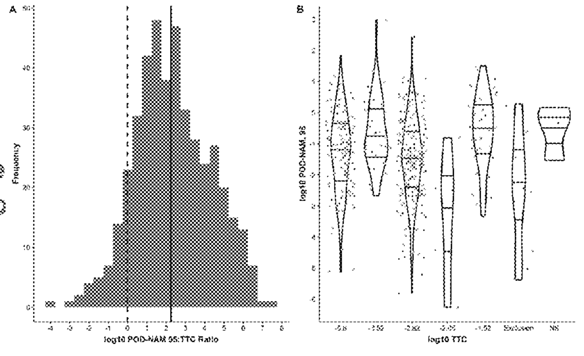
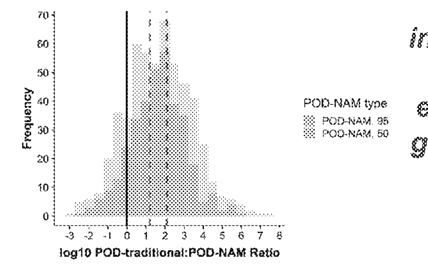


Figure 6. log_{10}POD ratio distribution

- log_{10}POD ratio is illustrated for the $POD_{NAM,95}$ and the $POD_{NAM,50}$.
- Using the more conservative (i.e., lower) $POD_{NAM,95}$, 48 of the 448 substances (10.7%) demonstrated a log_{10}POD ratio < 0 (to the left of the solid vertical line), whereas 92 of the 448 substances (20.5%) demonstrated a log10-POD ratio < 0 using the $POD_{NAM,50}$.
- The medians of the log10-POD ratio distributions are indicated by dashed lines for $POD_{NAM,95}$ and $POD_{NAM,50}$ as 2 and 1.2, respectively.



$POD_{NAM,95}$ includes interindividual variability in the in vitro to in vivo extrapolation process to a greater extent, and is more often a conservative estimate of $POD_{traditional}$

Figure 7. When the log_{10}POD ratio < 0, was it driven by a specific study type (as a surrogate for phenotypes)?

Condition	Dev/Repro is min POD	Dev/Repro is not min POD
log10-POD ratio,95 < 0	3	45
log10-POD ratio,95 > 0	41	359

Condition	Chronic is min POD	Chronic is not min POD
log10-POD ratio,95 < 0	28	20
log10-POD ratio,95 > 0	244	156

Based on a Fisher's exact test, when log_{10}POD ratio <0. it was not driven by a specific study type.

Figure 8. When the log_{10}POD ratio < 0, was it driven by a specific chemical features?

The enriched chemical structural features represented by ToxPrints for the log_{10}POD $ratio_{95}$ < 0 set.

ChemoType information		Appearance of the ToxPrint			Metrics			ChemoType information		Appearance of the ToxPrint			Metrics		
Label	ToxPrint	Total	POD ratio ≤ 0	POD ratio > 0	BA	OR	p-value	Label	ToxPrint	Total	POD ratio ≤ 0	POD ratio > 0	BA	OR	p-value
bond:P=O_phosphorus_oxo		18	12	6	0.62	22	7.4E-09	bond:P~N_generic		5	4	1	0.54	36	0.00055
bond:P=O_phosphate_thio		3	3	0	0.53	NA	0.0012	bond:C(=O)N_carbamate		20	6	14	0.54	3.9	0.014
bond:P~S_generic		27	13	14	0.62	10	3.5E-7	bond:CS_sulfide		53	15	38	0.61	4.3	0.00011

Based on a Fisher's exact test, chemical features associated with organophosphate pesticides and carbamates are more likely to drive a log_{10}POD ratio < 0.

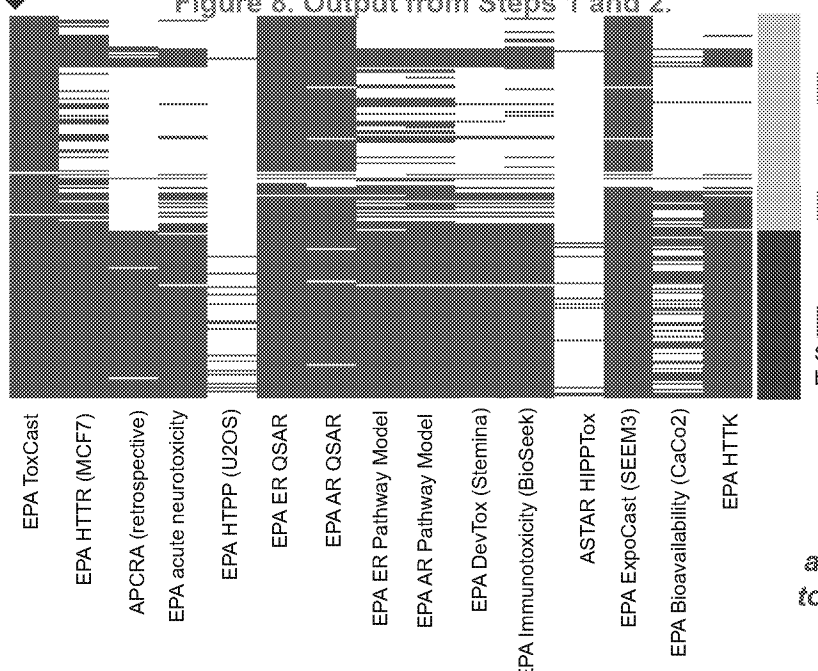
Part II: Prospective case study

- How well does a NAM-based approach perform in the prospective case?
- This prospective case study builds upon learnings from the retrospective case study, addressing questions including:
  - Can NAM-based POD estimates be improved using additional technologies or assumptions?
  - Are reasonable NAM-based POD estimates attainable for substances with limited *in vitro* bioactivity?
  - Can BER, and additional hazard flags, be used to select substances for *in vivo* screening?

Step 1 Identification of substances with:
- Limited hazard information and exposure potential
- Compatibility for currently available *in vitro* screening methodology

Step 2 Completion of a NAM battery for 200 substances within the substances identified
- Multiple *in vitro* platforms: ToxCast, high-throughput transcriptomics, high-throughput phenotypic profiling, Immunotoxicity assays, acute neurotoxicity assays, developmental toxicity assays, endocrine-relevant assays and models
- High-throughput toxicokinetic information for *in vitro* to *in vivo* extrapolation

Figure 8. Output from Steps 1 and 2.



In 2019, any gaps in this heatmap will be filled.

Scenario 1
Substance present on the EU, Canada, and/or US market, with a potential for consumer use and significant data gaps for systemic toxicity (105).

Scenario 2
Substance present on the EU, Canada, and/or US market, with known toxicity and potential interspecies differences (8).

Scenario 3
Substance selected from the retrospective case study, by sampling substances with varying log10POD ratios.

The BER (<10⁴) from Step 2, and hazard flags based on potential endocrine, developmental, neuro, and/or immuno-toxicity, will be used to advance ~20 substances to Step 3.

Step 3 Confirmatory 5-day *in vivo* testing based on BER and hazard flags
- Transcriptomics in liver
- Classical *in vivo* observations and toxicokinetics

Step 4 Further confirmation of a small subset from Step 3 in a 90-day subchronic study

Step 5 Evaluation
- Comparison of Step 2-4 data (if available), and any other traditional hazard information

Conclusions

- A major premise of this work is that the minimal concentration corresponding to *in vitro* bioactivity is likely to be a conservative threshold for any specific effects or toxicities that might be observed *in vivo*.
- BER may be a reasonable data-driven metric for prioritization that is tunable based on the amount of uncertainty in (1) the IVIVE that is included in development of the POD_{NAM} and (2) the exposure predictions, highlighting that for different screening applications differing amounts of uncertainty can be included in this workflow.
- The prospective case study furthers confidence, and identifies possible limitations, in NAM-based screening assessments.
- The collaborative, international consideration of these issues in screening level assessments demonstrates the current state-of-the-science and presents a transparent and adaptable basis for utilization of HTS information.